

STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/568,737
Source: IFW/P
Date Processed by STIC: 2/27/06

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.4.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER:

10/568,737

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped
 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence. (see item 11 below)
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section or use "chemically synthesized" as explanation. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32), also Sec. 1.823 of Sequence Rules
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWP

RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,737

TIME: 15:06:53

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02272006\J568737.raw

3 <110> APPLICANT: SHIRE BIOCHEM INC.
 5 <120> TITLE OF INVENTION: POLYPEPTIDES OF STREPTOCOCCUS PYOGENES
 7 <130> FILE REFERENCE: 51564-44
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/568,737
 C--> 10 <141> CURRENT FILING DATE: 2006-02-15
 12 <150> PRIOR APPLICATION NUMBER: US 60/495,094
 13 <151> PRIOR FILING DATE: 2003-08-15
 15 <160> NUMBER OF SEQ ID NOS: 44
 17 <170> SOFTWARE: PatentIn version 3.3
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 537
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Streptococcus pyogenes
 24 <400> SEQUENCE: 1
 25 atgtcacgta ttggtaataa agtaattact atgcctgcag gcgttgaatt aacaaataac 60
 27 aacaatgta ttactgttaa aggcctaaa ggcgaaactca ctcgtagatt caacaaaaat 120
 29 attgaaatca aagttgaagg gactgaaatc acagttgtac gtcctaacga ctcaaaagaa 180
 31 atgaaaacaa tccatggtag aaccctgtct aacttgaata acatgggtgt aggtgtttct 240
 33 gaagggtttca aaaaagatct tgaaatgaag ggtgtcgggt accgtgctca acttcaaggt 300
 35 actaaacttg tcctttcagt aggtaaatct caccaagacg aagttgaagc tccagaagga 360
 37 attactttca ctgttgctaa cccaacttca atctcagttg aaggaatcaa caaagaagtt 420
 39 gttggtcaaa cagctgctta catccgtagc ttgcgttcac cagagcctta caaaggcaaa 480
 41 gggatccgtt acgttggtga atacgtacgc cttaaagaag gtaaaacagg taaataa 537
 44 <210> SEQ ID NO: 2
 45 <211> LENGTH: 178
 46 <212> TYPE: PRT
 47 <213> ORGANISM: Streptococcus pyogenes
 49 <400> SEQUENCE: 2
 51 Met Ser Arg Ile Gly Asn Lys Val Ile Thr Met Pro Ala Gly Val Glu
 52 1 5 10 15
 55 Leu Thr Asn Asn Asn Val Ile Thr Val Lys Gly Pro Lys Gly Glu
 56 20 25 30
 59 Leu Thr Arg Glu Phe Asn Lys Asn Ile Glu Ile Lys Val Glu Gly Thr
 60 35 40 45
 63 Glu Ile Thr Val Val Arg Pro Asn Asp Ser Lys Glu Met Lys Thr Ile
 64 50 55 60
 67 His Gly Thr Thr Arg Ala Asn Leu Asn Asn Met Val Val Gly Val Ser
 68 65 70 75 80
 71 Glu Gly Phe Lys Lys Asp Leu Glu Met Lys Gly Val Gly Tyr Arg Ala
 72 85 90 95
 75 Gln Leu Gln Gly Thr Lys Leu Val Leu Ser Val Gly Lys Ser His Gln
 76 100 105 110
 79 Asp Glu Val Glu Ala Pro Glu Gly Ile Thr Phe Thr Val Ala Asn Pro

Does Not Comply
 Corrected Diskette Needed
 (Pg. 5)

RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,737

TIME: 15:06:53

Input Set : A:\PTO.RJ.txt

Output Set : N:\CRF4\02272006\J568737.raw

```

80          115          120          125
83 Thr Ser Ile Ser Val Glu Gly Ile Asn Lys Glu Val Val Gly Gln Thr
84          130          135          140
87 Ala Ala Tyr Ile Arg Ser Leu Arg Ser Pro Glu Pro Tyr Lys Gly Lys
88 145          150          155          160
91 Gly Ile Arg Tyr Val Gly Glu Tyr Val Arg Leu Lys Glu Gly Lys Thr
92          165          170          175
95 Gly Lys
99 <210> SEQ ID NO: 3
100 <211> LENGTH: 1269
101 <212> TYPE: DNA
102 <213> ORGANISM: Streptococcus pyogenes
104 <400> SEQUENCE: 3
105 atgtttcagt taagaaaaaa aatgacgcgc aaacaattag ccttggtgag tgctggagtg      60
107 ttgacctgtg tggttggtgg tagctacttg ataatgaacc atcaacaaca agaaattgtc      120
109 tctagtgtca acaaaagtaaa agccttaacc ataaaagaag ccatggaaca aggaaaagat      180
111 atcagcttga ccttagctgg cgaagtaaca gctaacaaca gcagcaaagt caaaatcgac      240
113 tcaagtaaag gagaagtcaa agaggctctt gttaaaaaag gcgatgttgt caaagtagga      300
115 caacccttgt tttagctatga aacgtcacag cggttaacgg ctcaaagttc agaatttgat      360
117 gttcaaacca aagccaatca gctccaagtt gctaaaacca atgcagcatt gaagtgggaa      420
119 acctacaatc gcaagggtcaa tgaaatcaac accctaaaat ctcgctacaa cactgcacca      480
121 gatgagagct tactagagca aattcgcagc gcagaagaca gtgtatccca agcactaagc      540
123 gatgccaaaa cagcagatag cgatgtcaaa accgctcaaa tcgaactcga taaagctaat      600
125 gctactgcca caacggaaaa aggtaaacta gagtatgaca ccgttaagtc agacaccgca      660
127 ggaaccattg ttagtctaaa tactgatttg ccaaataaat caaaatccaa aaaagaaaaat      720
129 gaaactttta tggaaattat cgacaaatca aaaatgtagg tcaaaggtaa cattagttaa      780
131 tttgaccgtg acaagttaaa aatcgggtcaa aaagtcgaag tgattgaccg caaagacaac      840
133 tctaaaaaat ggactggaaa agtaacccaa gttggcaacc tcaaagcaga ggaaaaaggc      900
135 caaggtcaag gccaaagggt caatgacca caagataatc caaaccaagc aaaattccct      960
137 tatgttattg aacttgacca atcagacaag cagccactca ttggctcaca cacctatgtt      1020
139 aatgtgctca acaatgttcc agaagctggc aagatcgtat tgaaagaaac ctttacaatg      1080
141 gcagaaaatg gaaaaaccta tgtgtggaaa gttgataaaa acaaggtcaa aaaacaagaa      1140
143 atcaagacta agcccttctc aaaaggttat gttgaggtaa caagtggctt gactatgcaa      1200
145 gataagattg ctacgccgct tcctggcatg aaagacggtg tggaggtagg aagtattgtt      1260
147 aaaccttaa
150 <210> SEQ ID NO: 4
151 <211> LENGTH: 422
152 <212> TYPE: PRT
153 <213> ORGANISM: Streptococcus pyogenes
155 <400> SEQUENCE: 4
157 Met Phe Gln Leu Arg Lys Lys Met Thr Arg Lys Gln Leu Ala Leu Leu
158 1          5          10          15
161 Ser Ala Gly Val Leu Thr Cys Val Val Gly Gly Ser Tyr Leu Ile Met
162          20          25          30
165 Asn His Gln Gln Gln Glu Ile Val Ser Ser Val Asn Lys Val Lys Ala
166          35          40          45
169 Leu Thr Ile Lys Glu Ala Met Glu Gln Gly Lys Asp Ile Ser Leu Thr
170          50          55          60
173 Leu Ala Gly Glu Val Thr Ala Asn Asn Ser Ser Lys Val Lys Ile Asp

```

RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,737

TIME: 15:06:53

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02272006\J568737.raw

```

174 65          70          75          80
177 Ser Ser Lys Gly Glu Val Lys Glu Val Phe Val Lys Lys Gly Asp Val
178          85          90          95
181 Val Lys Val Gly Gln Pro Leu Phe Ser Tyr Glu Thr Ser Gln Arg Leu
182          100          105          110
185 Thr Ala Gln Ser Ser Glu Phe Asp Val Gln Thr Lys Ala Asn Gln Leu
186          115          120          125
189 Gln Val Ala Lys Thr Asn Ala Ala Leu Lys Trp Glu Thr Tyr Asn Arg
190          130          135          140
193 Lys Val Asn Glu Ile Asn Thr Leu Lys Ser Arg Tyr Asn Thr Ala Pro
194 145          150          155          160
197 Asp Glu Ser Leu Leu Glu Gln Ile Arg Ser Ala Glu Asp Ser Val Ser
198          165          170          175
201 Gln Ala Leu Ser Asp Ala Lys Thr Ala Asp Ser Asp Val Lys Thr Ala
202          180          185          190
205 Gln Ile Glu Leu Asp Lys Ala Asn Ala Thr Ala Thr Thr Glu Lys Gly
206          195          200          205
209 Lys Leu Glu Tyr Asp Thr Val Lys Ser Asp Thr Ala Gly Thr Ile Val
210          210          215          220
213 Ser Leu Asn Thr Asp Leu Pro Asn Gln Ser Lys Ser Lys Lys Glu Asn
214 225          230          235          240
217 Glu Thr Phe Met Glu Ile Ile Asp Lys Ser Lys Met Leu Val Lys Gly
218          245          250          255
221 Asn Ile Ser Glu Phe Asp Arg Asp Lys Leu Lys Ile Gly Gln Lys Val
222          260          265          270
225 Glu Val Ile Asp Arg Lys Asp Asn Ser Lys Lys Trp Thr Gly Lys Val
226          275          280          285
229 Thr Gln Val Gly Asn Leu Lys Ala Glu Glu Lys Gly Gln Gly Gln Gly
230          290          295          300
233 Gln Gly Gly Asn Asp Gln Gln Asp Asn Pro Asn Gln Ala Lys Phe Pro
234 305          310          315          320
237 Tyr Val Ile Glu Leu Asp Gln Ser Asp Lys Gln Pro Leu Ile Gly Ser
238          325          330          335
241 His Thr Tyr Val Asn Val Leu Asn Asn Val Pro Glu Ala Gly Lys Ile
242          340          345          350
245 Val Leu Lys Glu Thr Phe Thr Met Ala Glu Asn Gly Lys Thr Tyr Val
246          355          360          365
249 Trp Lys Val Asp Lys Asn Lys Val Lys Lys Gln Glu Ile Lys Thr Lys
250          370          375          380
253 Pro Phe Ser Lys Gly Tyr Val Glu Val Thr Ser Gly Leu Thr Met Gln
254 385          390          395          400
257 Asp Lys Ile Ala Gln Pro Leu Pro Gly Met Lys Asp Gly Met Glu Val
258          405          410          415
261 Gly Ser Ile Val Lys Pro
262          420
265 <210> SEQ ID NO: 5
266 <211> LENGTH: 885
267 <212> TYPE: DNA
268 <213> ORGANISM: Streptococcus pyogenes

```

RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,737

TIME: 15:06:53

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02272006\J568737.raw

270 <400> SEQUENCE: 5

```

271 atgataaaac gatgtaaagg aattgggtcta gccttaatgg ccttcttttt ggtagcttgt      60
273 gtgaatcagc accctaaaac ggctaaagag actgaacagc agagaattgt agccacttcg      120
275 gttgctgtgg ttgatattctg tgaccgttta aatttagacc tcggtggggg ttgtgatagt      180
277 aaattatata ccttccttaa acgctatgat gctgttaagc gtgtggggtt acccatgaat      240
279 cctgatattag agttgattgc ttctttgaaa ccaacttga ttttgagtcc caattcttta      300
281 caagaagatt tggaacccaa gtatcaaaaa ttggatactg agtatggttt tttgaactta      360
283 cgaagtgttg agggcatgta ccagtcatt gatgatttag ggaacctttt ccaacgtcaa      420
285 caagaagcaa aagaattgag ccagcaatac caggactatt atcgtgcttt ccaagctaaa      480
287 cgtaaggagg agaaaaagcc taaagtgcct attcttatgg gcttgccagg tagttatttg      540
289 gtggcgacga accaatctta tgtagggaat cttttggact tggcagggtg tgagaatggt      600
291 tatcagtcag atgagaaaga atttctatca gctaactctg aagacatgct ggctaaggag      660
293 cctgacttga ttttacgaac agctcatgcc attccagaca aggtaaaagt gatgtttgac      720
295 aaagaatttg ctgaaaatga tatttgaaa cattttacgg cagtcaagga agggaaagtc      780
297 tatgatttgg acaataccct gtttggcatg agtgctaaat tgaactaccc agaagccttg      840
299 gacaccttaa cacagctttt tgaccacgtg ggagatcatc cgtaa                        885

```

302 <210> SEQ ID NO: 6

303 <211> LENGTH: 294

304 <212> TYPE: PRT

305 <213> ORGANISM: Streptococcus pyogenes

307 <400> SEQUENCE: 6

```

309 Met Ile Lys Arg Cys Lys Gly Ile Gly Leu Ala Leu Met Ala Phe Phe
310 1          5          10          15
313 Leu Val Ala Cys Val Asn Gln His Pro Lys Thr Ala Lys Glu Thr Glu
314          20          25          30
317 Gln Gln Arg Ile Val Ala Thr Ser Val Ala Val Val Asp Ile Cys Asp
318          35          40          45
321 Arg Leu Asn Leu Asp Leu Val Gly Val Cys Asp Ser Lys Leu Tyr Thr
322          50          55          60
325 Leu Pro Lys Arg Tyr Asp Ala Val Lys Arg Val Gly Leu Pro Met Asn
326 65          70          75          80
329 Pro Asp Ile Glu Leu Ile Ala Ser Leu Lys Pro Thr Trp Ile Leu Ser
330          85          90          95
333 Pro Asn Ser Leu Gln Glu Asp Leu Glu Pro Lys Tyr Gln Lys Leu Asp
334          100         105         110
337 Thr Glu Tyr Gly Phe Leu Asn Leu Arg Ser Val Glu Gly Met Tyr Gln
338          115         120         125
341 Ser Ile Asp Asp Leu Gly Asn Leu Phe Gln Arg Gln Gln Glu Ala Lys
342          130         135         140
345 Glu Leu Arg Gln Gln Tyr Gln Asp Tyr Tyr Arg Ala Phe Gln Ala Lys
346 145         150         155         160
349 Arg Lys Gly Lys Lys Lys Pro Lys Val Leu Ile Leu Met Gly Leu Pro
350          165         170         175
353 Gly Ser Tyr Leu Val Ala Thr Asn Gln Ser Tyr Val Gly Asn Leu Leu
354          180         185         190
357 Asp Leu Ala Gly Gly Glu Asn Val Tyr Gln Ser Asp Glu Lys Glu Phe
358          195         200         205
361 Leu Ser Ala Asn Pro Glu Asp Met Leu Ala Lys Glu Pro Asp Leu Ile
362          210         215         220

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/568,737

DATE: 02/27/2006
TIME: 15:06:53

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02272006\J568737.raw

365 Leu Arg Thr Ala His Ala Ile Pro Asp Lys Val Lys Val Met Phe Asp
366 225 230 235 240
369 Lys Glu Phe Ala Glu Asn Asp Ile Trp Lys His Phe Thr Ala Val Lys
370 245 250 255
373 Glu Gly Lys Val Tyr Asp Leu Asp Asn Thr Leu Phe Gly Met Ser Ala
374 260 265 270
377 Lys Leu Asn Tyr Pro Glu Ala Leu Asp Thr Leu Thr Gln Leu Phe Asp
378 275 280 285
381 His Val Gly Asp His Pro
382 290
385 <210> SEQ ID NO: 7
386 <211> LENGTH: 34
387 <212> TYPE: DNA
388 <213> ORGANISM: Primer *Invalid response*
390 <400> SEQUENCE: 7
391 gagaaaatac atatgtcacg tattggtaat aaag
394 <210> SEQ ID NO: 8
395 <211> LENGTH: 29
396 <212> TYPE: DNA
397 <213> ORGANISM: Primer *Invalid Response*
399 <400> SEQUENCE: 8
400 ccctcgagtt atttacctgt tttaccttc
403 <210> SEQ ID NO: 9
404 <211> LENGTH: 31
405 <212> TYPE: DNA
406 <213> ORGANISM: Primer *Same error*
408 <400> SEQUENCE: 9
409 aaggatccca tggtcacgtat tggtaataaa g
412 <210> SEQ ID NO: 10
413 <211> LENGTH: 38
414 <212> TYPE: DNA
415 <213> ORGANISM: Primer *Same error*
417 <400> SEQUENCE: 10
418 actagtcgac ttatttacct gttttacctt ctttaagg
421 <210> SEQ ID NO: 11
422 <211> LENGTH: 35
423 <212> TYPE: DNA
424 <213> ORGANISM: Primer *Same error*
426 <400> SEQUENCE: 11
427 ccttacaag gcaaaggcat ccgttacgtt ggtga
430 <210> SEQ ID NO: 12
431 <211> LENGTH: 35
432 <212> TYPE: DNA
433 <213> ORGANISM: Primer *Same error*
435 <400> SEQUENCE: 12
436 tcaccaacgt aacggatgcc tttgcctttg taagg
439 <210> SEQ ID NO: 13
440 <211> LENGTH: 30
441 <212> TYPE: DNA

34

29

31

38

35

35

The type of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/568,737

DATE: 02/27/2006

TIME: 15:06:54

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02272006\J568737.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date